

Qualitätssicherung bei qualitativ inhaltsanalytischen Verfahren in großen Forschungsgruppen: Herausforderungen und Möglichkeiten in der Forschungspraxis am Beispiel der perLen-Studie

Vasarik Staub, Katriina; Stebler, Rita; Reusser, Kurt; Galle, Marco

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Vasarik Staub, K., Stebler, R., Reusser, K., & Galle, M. (2019). Qualitätssicherung bei qualitativ inhaltsanalytischen Verfahren in großen Forschungsgruppen: Herausforderungen und Möglichkeiten in der Forschungspraxis am Beispiel der perLen-Studie. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), 1-20. <https://doi.org/10.17169/fqs-20.3.3391>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see:
<https://creativecommons.org/licenses/by/4.0>

Qualitätssicherung bei qualitativ inhaltsanalytischen Verfahren in großen Forschungsgruppen: Herausforderungen und Möglichkeiten in der Forschungspraxis am Beispiel der perLen-Studie

Katriina Vasarik Staub, Marco Galle, Rita Stebler & Kurt Reusser

Keywords:

qualitative
Inhaltsanalyse;
Qualitätssicherung;
große Forschungsgruppen;
multimethodisches
Längsschnittdesign;
Intercoder-Übereinstimmung

Zusammenfassung: Strategien der Qualitätssicherung sind ein vielversprechender und zunehmend beliebter Ansatz zur Qualitätskontrolle in qualitativen Studien. Besonders wichtig sind sie in großen Forschungsgruppen. Am Beispiel der multimethodischen Längsschnittstudie [Personalisierte Lernkonzepte in heterogenen Lerngruppen – perLen](#) werden Herausforderungen und Möglichkeiten bei der Anwendung von Strategien der Qualitätssicherung in qualitativ inhaltsanalytischen Teilstudien untersucht und reflektiert. Dabei zeigt sich, dass günstige Wirkungen auf die Qualität des gesamten Forschungsprozesses und dadurch auf die Ergebnisse zu erwarten sind, wenn Kriterien und Strategien der methodenspezifischen und der methodenübergreifenden Qualitätssicherung kombiniert, Erhebungs- und Auswertungsverfahren aufeinander abgestimmt, ein einheitliches Verfahrens- und Dokumentationssystem etabliert, alle Mitwirkenden mit Instrumenten und Verfahren vertraut gemacht und Qualitätskontrollen nicht nur methodenspezifisch und prozessorientiert, sondern in jedem einzelnen Forschungsschritt integrativ durchgeführt werden. Die hier mit Blick auf die qualitative Inhaltsanalyse beschriebenen Erkenntnisse können in ihren Grundzügen auf Untersuchungen mit anderen qualitativen Methoden übertragen werden und dazu beitragen, die Qualität innerhalb und zwischen den Erhebungszeitpunkten in qualitativen Längsschnittstudien zu sichern und dadurch auch die Aussagekraft multimethodischer Studien zu erhöhen.

Inhaltsverzeichnis

- [1. Einleitung](#)
 - [2. Qualitätssicherung in qualitativen Forschungsprozessen](#)
 - [3. Qualitätssicherung in der Forschungspraxis – ein Beispiel](#)
 - [3.1 Leitfadeninterviews in der perLen-Studie](#)
 - [3.2 Qualitätssicherung der Interviewdaten: Herausforderungen und Lösungen](#)
 - [3.3 Konsequenzen](#)
 - [4. Maßnahmen zur Qualitätssicherung für große Forschungsgruppen](#)
- [Danksagung](#)
- [Literatur](#)
- [Zu den Autorinnen und Autoren](#)
- [Zitation](#)

1. Einleitung

Qualitätssicherung ist in der qualitativen Forschung seit Langem ein Thema (LUDWIG 2011; STRÜBING, HIRSCHAUER, AYASS, KRÄHNKE & SCHEFFER 2018).¹ In den 1990er Jahren wurden im Qualitätsdiskurs vorwiegend die positivistischen bzw. "klassischen" Gütekriterien diskutiert, im Vordergrund stand die Legitimierung der qualitativen Forschung (HELSPER, KELLE & KOLLER 2016). Heute erscheinen forschungspraktische Fragen der Qualitätssicherung (u.a. Welche Qualitätsziele haben Forschungsprojekte? Was trägt zur Glaubwürdigkeit und Nachvollziehbarkeit der Vorgehensweisen und Ergebnisse nach außen bei?) zentral (FLICK 2017). Mit dem vorliegenden Beitrag betrachten und diskutieren wir Qualitätssicherung am Beispiel der qualitativen Inhaltsanalyse² in einem großen, längsschnittlich angelegten Forschungsprojekt mit multimethodischem Design (STEBLER, PAULI & REUSSER 2018). Das Ziel unseres Beitrags besteht nicht darin, grundlegend neue Erkenntnisse zur Qualitätsdebatte beizutragen. Wir sind vielmehr daran interessiert aufzuzeigen, wie sich etablierte, methodenspezifische und ergebnisorientierte Qualitätssicherungsmaßnahmen (KUCKARTZ 2018) mit methodenübergreifenden und prozessorientierten Qualitätssicherungsmaßnahmen (STRÜBING et al. 2018) kombinieren lassen. Anhand eines Beispiels aus der Forschungspraxis erörtern wir, dass die qualitative Inhaltsanalyse aus der Perspektive der Qualitätssicherung nicht isoliert betrachtet werden kann, sondern dass insbesondere im Kontext multimethodischer und längsschnittlich angelegter Forschungsprojekte mit großen Forschungsgruppen eine komplexe und strategische Qualitätssicherung mit Wirkung auf den ganzen Forschungsprozess an Bedeutung gewinnt. Darüber hinaus argumentieren wir auf Grundlage unserer Forschungserfahrungen, dass das Zusammenwirken von Datenerhebung und Datenauswertung im Rahmen der Qualitätssicherung als Gegenstand wissenschaftlicher Projektdesigns und Schulung von Forscher_innen stärker berücksichtigt werden sollte. [1]

In der Bildungs- und Unterrichtsforschung werden empirische Fragestellungen zunehmend multimethodisch angegangen (z.B. BRAUN, WEISS & KIEL 2018). Dabei werden in großen Forschungsprojekten kleine, qualitativ ausgerichtete Teilstudien durchgeführt, um zentrale Aspekte der Fragestellung inhaltlich zu vertiefen (z.B. PILZ 2018; REINOLD 2016). Diese Teilstudien werden meistens querschnittlich angelegt (z.B. BRAUN et al. 2018) und von einzelnen Forscher_innen durchgeführt (z.B. HÄBIG 2018). Große multimethodische Forschungsprojekte mit längsschnittlich konzipierten, qualitativen Teilstudien sind bisher eher die Ausnahme (z.B. BOHL & WACKER 2016), u.a. weil solche Studien personalintensiv und bezüglich der Komplexität der Fragestellungen, des Umfangs der Datenerhebungen sowie der Qualitätssicherung äußerst anspruchsvoll sind. Demgegenüber gewinnen multimethodische Designs in der Sozialforschung immer mehr an Bedeutung (HELSPER 2016; KUCKARTZ 2018). Durch die Verbindung quantitativer und qualitativer Zugänge sollen

1 Siehe auch die [FQS-Debatte zu Qualitätsstandards qualitativer Sozialforschung](#).

2 Siehe für einen Überblick über die Methode und ihre Varianten SCHREIER (2014) und STAMANN, JANSSEN und SCHREIER (2016).

generalisierbare, am Phänomen ausgerichtete und extern hinreichend valide Ergebnisse erzeugt werden. Es ist also anzunehmen, dass Forscher_innen sich in Bezug auf die Qualitätssicherung auf zunehmend komplexere Forschungsprozesse einstellen müssen. [2]

Große längsschnittlich angelegte Forschungsprojekte sind in der Regel durch multiple Datenerhebungs- und Analyseprozesse zu mehreren Forschungsfragen gekennzeichnet (ASBRAND, PFAFF & BOHNSACK 2013). Daraus ergeben sich vielschichtige Herausforderungen für den Forschungsprozess. Neben Fragebogen- und Leistungsdaten werden häufig auch umfangreiche Sammlungen von Text-, Bild- und Tondokumenten als Daten gewonnen. Je nach Fragestellung wird ein breites methodologisches Instrumentarium angewendet. Weil Forscher_innen mit Längsschnittdesigns stets darauf zielen, Wandel oder Entwicklungsverläufe zu untersuchen (a.a.O.), muss hier Qualität über die Erhebungszeitpunkte hinweg gewährleistet werden. Das heißt, die aufeinanderfolgenden Erhebungen müssen inhaltlich und methodisch aufeinander bezogen und die Daten konsistent aufbereitet werden. Besonders in den qualitativen Projektteilen erfolgen die Vorbereitungen (u.a. Instrumentenentwicklung, Schulung der Interviewer_innen), die Datenerhebungen (u.a. Interviews), das Datenmanagement (u.a. Transkription, Archivierung) und die Auswertungen (u.a. qualitative Inhaltsanalyse) iterativ und phasenweise parallel (STRÜBING 2018). Dies erfordert große Forschungsgruppen, die erfahrungsgemäß mit einer hohen personellen Fluktuation während der Projektlaufzeit einhergehen. Damit unter diesen Vorzeichen auch die qualitativen Studienteile wissenschaftlichen Ansprüchen genügen, ist eine umfassende, prozessorientierte Qualitätssicherung (FLICK 2018) unerlässlich. Bislang ist wenig darüber bekannt, wie sich unterschiedliche methodenspezifische und methodenübergreifende Maßnahmen in solchen Situationen zielbezogen verbinden lassen. Es erscheint uns daher sinnvoll, über praxistaugliche Kombinationen methodenspezifischer und -übergreifender sowie produkt- und prozessorientierter Kriterien und Strategien der Qualitätssicherung nachzudenken. [3]

Unsere Ausführungen sind an zwei Fragen ausgerichtet: 1. Mit welchen Maßnahmen können Forscher_innen im Forschungsprozess die Qualität von qualitativ inhaltsanalytischen Zugängen günstig beeinflussen? 2. Welche Möglichkeiten gibt es, um die Qualität in den verschiedenen, parallel ablaufenden und längsschnittlich aufeinander bezogenen Datenerhebungs- und Auswertungsphasen zu sichern? Im Folgenden werden die Qualitätssicherung in der qualitativen Forschung beschrieben und Bezüge zur qualitativen Inhaltsanalyse hergestellt (Abschnitt 2). Danach werden am Beispiel des Projekts [Personalisierte Lernkonzepte in heterogenen Lerngruppen – perLen](#) (STEBLER et al. 2018) Herausforderungen und Möglichkeiten der Qualitätssicherung bei qualitativ inhaltsanalytischen Verfahren in großen Längsschnittstudien dargestellt und mit Bezug auf Aspekte des Qualitätsdiskurses reflektiert (Abschnitt 3). Wir schließen den Beitrag mit einer zusammenfassenden Erörterung von Maßnahmen zur Qualitätssicherung im aktuellen Qualitätsdiskurs mit Blick auf große Forschungsgruppen (Abschnitt 4). [4]

2. Qualitätssicherung in qualitativen Forschungsprozessen

Innerhalb des qualitativen Forschungsparadigmas fehlt bisher eine verbindliche Definition von *Qualität*. Dies liegt u.a. an den unterschiedlichen Ansätzen, Schulen und Richtungen (BERGMAN et al. 2010), den sprachraumspezifischen Sichtweisen und den unterschiedlichen Anwendungsfeldern qualitativer Forschung (FLICK 2005). Zudem wurde das Verständnis dessen, was Qualität in der qualitativen Forschung ausmacht, in den letzten Jahrzehnten ausdifferenziert (z.B. DENZIN & LINCOLN 2005; FLICK 2007a). Der Qualitätsdiskurs wird vorwiegend disziplinintern (u.a. Soziologie: SCHÜTZE 2005; Psychologie: GERGEN 2014; WERTZ 2011; Gesundheitswesen: ELO et al. 2014; LI & SEALE 2007) oder methodenspezifisch (BARBOUR 2014; BOHNSACK & KRÜGER 2005; DENZIN & LINCOLN 2005; FLICK 2014, 2018; LÜDERS 2013; STEINKE 2013) geführt. "Quality is elusive, hard to pre-specify, but we often feel we know it when we see it. In this respect research is like art rather than science" (SEALE 2002, S.102). In der Forschungspraxis besteht jedoch die Notwendigkeit, diese Kunst begrifflich zu fassen, sodass Forschungsprozesse und -ergebnisse auch für Außenstehende nachvollziehbar sind (SCHÜTZE 2005). [5]

Gemeinsam ist vielen Ansätzen zur Qualitätssicherung, dass sowohl erhebungs- und auswertungsbezogene als auch ergebnis- und prozessorientierte Kriterien mitgedacht werden (BERGMAN & COXON 2005; BERGMAN et al. 2010; KRÜGER 2010; MAYRING 2016; SEALE 1999; STEINKE 2013; STRÜBING et al. 2018). Allerdings besteht ein disziplinübergreifender Konsens, dass die Qualität qualitativer Forschung mittels positivistischer (quantitativer) Gütekriterien (u.a. Objektivität, Reliabilität, Validität) nicht hinreichend beschrieben werden kann (BARBOUR 2014; FLICK 2018; FROST 2014). Um die Qualität der Erkenntnisse unterschiedlicher Methoden zu überprüfen, werden eher generelle Kriterien vorgeschlagen: "Diese Kriterien dienen als Zielvorgaben und Prüfsteine einer beliebigen angewandten Forschungsmethode, an denen der Grad der Wissenschaftlichkeit dieser Methode gemessen werden kann" (LAMNEK & KRELL 2016, S.141). In der Literatur findet man ein breites Kriterienspektrum, das von der Reformulierung der positivistischen Gütekriterien (BERGMAN et al. 2010) bis hin zu alternativen, methodenspezifischen Qualitätsmerkmalen des wissenschaftlichen Outputs (FLICK 2007b) reicht. [6]

So rückt FLICK (2018) beispielsweise den Forschungsprozess in den Vordergrund. Die Logik hinter den Bestrebungen einer *prozessorientierten Qualitätssicherung* lässt sich mit dem Verständnis von Qualität als Perfektion oder Konsistenz (HARVEY & GREEN 2000) beschreiben. Hier können die Endprodukte (Daten, Ergebnisse etc.) als qualitativ hochwertig gelten, wenn im Prozess qualitätsvolles Handeln erfolgt. Im Gegensatz dazu ist im Sinne der *ergebnisorientierten Qualitätssicherung* (a.a.O., siehe auch KIENER & SCHANNE 2001) ein Endprodukt qualitativ hochwertig, wenn es bestimmte Qualitätsstandards bzw. Gütekriterien (BARBOUR 2014; DÖRING & BORTZ 2016; FROST 2014; LAMNEK 2010; SEALE 2002) erfüllt (Qualität als Ausnahme). [7]

Unserer Meinung nach ist ein Vorteil von Strategien der Qualitätssicherung gegenüber Gütekriterien deren breites Anwendungsspektrum. Während sich beispielsweise die positivistischen Gütekriterien stark auf die Datenerhebung oder die Datenauswertung beziehen, lassen sich Strategien der Qualitätssicherung mehrheitlich auf den ganzen Forschungsprozess, von der Entwicklung des Forschungsdesigns bis zur Publikation der Ergebnisse, anwenden (BERGMAN & COXON 2005). Als *Strategien der Qualitätssicherung* gelten somit studienspezifische Maßnahmen, die mit dem Ziel angewendet werden, die Qualität des ganzen Forschungsprozesses zu sichern. Beispiele für Strategien sind Distanzierung (die Forscher_innen distanzieren sich bewusst von den eigenen Vorannahmen; BOHNSACK & KRÜGER 2005), Indikation (Abstimmen der Forschungsmethoden auf die Forschungsfrage sowie Reflexion von Forschungsschritten und -methoden; FLICK 2014), analytische Induktion (vorläufige Interpretationen werden an abweichenden Fällen überprüft; FLICK 2007b) und die Triangulation von Daten, Methoden, Theorien und Forscher_innen zwecks systematischer Erweiterung der Erkenntnismöglichkeiten (FLICK 2011, 2018). Auch die Verfahrensbeschreibungen für qualitative Methoden werden zu den Strategien der Qualitätssicherung gezählt (BERGMAN & COXON 2005). [8]

Je genauer ein Verfahren beschrieben wird, desto wahrscheinlicher scheint uns ein hoher Standardisierungsgrad, eine weitgehende Nachvollziehbarkeit für Außenstehende sowie eine gute Vergleichbarkeit der Ergebnisse. Qualität wird durch Transparenz, Genauigkeit und Stringenz prozessübergreifend gewährleistet (SEALE 2002). Leitlinien – u.a. bezogen auf die Fragestellung, das Sampling, den Umgang mit der Subjektivität der Forscher_innen, die Erhebung und die Analyse der Daten sowie die Interpretation der Ergebnisse (BERGMAN & COXON 2005; GERGEN 2014; SEALE 1999) – dienen dazu, Forschungsprozesse zu kontrollieren und zu bewerten. Die Wirksamkeit von Strategien hängt davon ab, wie konsequent sie verfolgt werden (BERGMAN & COXON 2005). Die Nutzung methodenspezifischer Kriterien der Qualitätssicherung, wie z.B. Intercoder-Übereinstimmung bei qualitativer Inhaltsanalyse (KUCKARTZ 2018), kann dazu beitragen, die Wirksamkeit der Strategien zu überprüfen. [9]

In Tabelle 1 ist eine Auswahl methodenspezifischer und methodenübergreifender Maßnahmen der Qualitätssicherung zusammengestellt, die sich entweder vorwiegend auf das Produkt oder den Prozess der Forschung beziehen. Es handelt sich dabei um einen Versuch, Kriterien und Strategien der Qualitätssicherung in Verbindung mit der qualitativen Inhaltsanalyse, welche als einzelne Verfahren bekannt und in der Forschungspraxis verankert sind zu ordnen. Dabei muss beachtet werden, dass die Verwendung der Begriffe "Kriterium" und "Strategie" sowohl für die qualitative Inhaltsanalyse als auch für weitere Verfahren u.a. von der Situation bzw. der Perspektive der Betrachter_innen abhängt. Ein Beispiel: Wenden Forscher_innen im Rahmen einer qualitativen Inhaltsanalyse das Konsensverfahren (HOPF & SCHMIDT 1993) an, wird üblicherweise von einer Strategie gesprochen. Externe Gutachter_innen hingegen können in der Regel nur feststellen, ob das Verfahren

eingesetzt wurde bzw. das Kriterium erfüllt ist. Aus diesem Grund wird in diesem Beitrag "Maßnahmen der Qualitätssicherung" als Oberbegriff für Kriterien und Strategien verwendet. Aus unserer Sicht hilft es mit Blick auf die Qualität einer Studie, wenn Forscher_innen Maßnahmen der Qualitätssicherung in den Studiendesigns festhalten und dadurch Weichenstellungen für eine integrative, projektspezifische Anwendung von methodenspezifischen und methodenübergreifenden Strategien (prozessorientiert) und Kriterien (produktorientiert) vornehmen.

	Methodenspezifische Qualitätssicherung am Beispiel der qualitativen Inhaltsanalyse	Methodenübergreifende Qualitätssicherung
Produktorientierte Qualitätssicherung	Intercoder-Übereinstimmung Konsensverfahren	Transkriptionsregeln Anonymisierung von Daten Kommunikative Validierung Triangulation Relevanz der Fragestellung und Theorie Kohärenz der Theorie
Prozessorientierte Qualitätssicherung	Training der Interviewer_innen je nach Art des Interviews Regeln zum Kameraeinsatz bei videogestützten Beobachtungen Computergestützte Datenanalyse Forschungstagebuch für intersubjektive Nachvollziehbarkeit Training der Auswerter_innen je nach Art der Datenanalyse	Forschungs- und Verfahrensdokumentation Regeln zur Archivierung, Speicherung, Nennung der Daten Datenschutz Regelgeleitetheit Intersubjektive Nachvollziehbarkeit Reflektierte Subjektivität

Tabelle 1: Maßnahmen der Qualitätssicherung [10]

Vor dem Hintergrund der beschriebenen komplexen Herausforderungen in großen, längsschnittlich angelegten Forschungsgruppen stellt sich die Frage nach dem *Zusammenspiel* der einzelnen Maßnahmen der Qualitätssicherung. Für die Intercoder-Übereinstimmung (z.B. KUCKARTZ 2018), die neben dem oben erwähnten Konsensverfahren (HOPF & SCHMIDT 1993) eine wichtige Maßnahme der Qualitätssicherung ist, empfiehlt sich der Einsatz einer entsprechenden Forschungssoftware. Auch das Konsensverfahren kann mit einem Computerprogramm schnell und effektiv durchgeführt werden (KUCKARTZ 2018). Mittels Software können beispielsweise Codierungsarbeiten

einfach dokumentiert und nachvollzogen werden, was die Transparenz der qualitativen Datenanalyse erhöht (KELLE & LAURIE 1995). Zu bedenken ist jedoch, dass der Computereinsatz an sich keine Qualität erzeugt: Er bietet Möglichkeiten einer technisch hochwertigen Verfahrens- und Analysedokumentation (KRIPPENDORFF 2013; MAYRING 2015), kann aber keine komplexe automatische Codierung im Sinne einer sachlich fundierten Auswertungsarbeit durchführen (KUCKARTZ 2018). Umgekehrt liegt darin der Mehrwert des Einsatzes einer Forschungssoftware für den Forschungsprozess: Eine regelgeleitete, inhaltsbezogene Anwendung von Analyseprogrammen kann zur prozessbezogenen Qualitätssicherung beitragen. Dies zeigen auch Erfahrungsberichte und Best-Practice-Beispiele zu computergestützt durchgeführten qualitativen Datenanalysen in großen Forschungsgruppen (BOURDON 2002; GIBBS, FRIESE & MANGABEIRA 2002; KUCKARTZ 2018). Das Fazit ist positiv: "The computer is also a powerful part of the context in social terms and this needs to be considered in deciding how to arrange groups for collaborative working" (FORD, OBERSKI & HIGGINS 2000, S.12). Mit Analyseprogrammen ist es möglich, die Aufbereitung und Auswertung von Daten in Forschungsgruppen zu koordinieren und arbeitsteilig durchzuführen. Grundlegend ist dabei eine sachdienliche und transparente Arbeits- und Datenorganisation (BOURDON 2002; EVERS et al. 2011; FORD et al. 2000; GIBBS et al. 2002; KUCKARTZ 2018). [11]

Bei methodenübergreifenden und prozessorientierten Strategien der Qualitätssicherung ist zentral, dass sich alle Forscher_innen für die Qualität der Prozesse verantwortlich fühlen, in diesem Sinne gemeinsam handeln und somit gemeinsam Qualität herstellen (FLICK 2014). Hier stellt sich die Frage, wie garantiert werden kann, dass die Strategien der Qualitätssicherung tatsächlich von allen Forscher_innen und in jedem einzelnen Forschungsschritt eingesetzt werden. Im Folgenden zeigen wir Möglichkeiten anhand eines forschungspraktischen Beispiels auf, basierend auf einer qualitativ inhaltsanalytischen Teilstudie. [12]

3. Qualitätssicherung in der Forschungspraxis – ein Beispiel

Personalisierte Lernkonzepte als Antwort auf heterogene Lerngruppen haben sich an vielen Schweizer Schulen aus der Praxis heraus entwickelt. In der multimethodischen Längsschnittstudie [Personalisierte Lernkonzepte in heterogenen Lerngruppen – perLen](#)³ wurde diese Praxis untersucht. Konkret im Fokus standen der Unterricht und die Lernkulturen in solchen Schulen, die sich dezidiert an "personalisierten Lernkonzepten" orientieren. Das Profil der teilnehmenden Schulen war in hohem Maße auf individualisierende und kooperative Formen des Lernens und adaptive Lernunterstützung ausgerichtet. Die Unterrichtsorganisation an diesen Schulen ist teils deutlich von klassischen Mustern abgewichen. Im Rahmen der Studie wurden 65 öffentliche und private Schulen aus der Deutschschweiz, in denen Unterrichts- und Schulentwicklung in Richtung personalisiertes Lernen betrieben wird, untersucht. Es wurden u.a.

3 Laufzeit: 2013-2015, die Studie wurde von der Stiftung Mercator Schweiz gefördert.

Online-Befragungen, Leistungstests, videogestützte Unterrichtsbeobachtungen und Leitfadeninterviews durchgeführt. Dabei interessierten in erster Linie Prozessqualitäten sowie die Unterrichtsformen, die damit zusammenhängenden Rollen und Anforderungen an die Lehrpersonen sowie die fachlichen und überfachlichen Lernerträge (STEBLER et al. 2018). [13]

3.1 Leitfadeninterviews in der perLen-Studie

In drei Phasen wurden Daten in einem komplexen mehrperspektivischen, längsschnittlichen und multimethodischen Projektdesign (BRYMAN 2006) erhoben. Zwölf der 65 Schulen wurden als sogenannte Fallschulen behandelt. In diesen wurden zusätzlich zu standardisierten Erhebungen und Leistungstests auch qualitative Methoden, u.a. Leitfadeninterviews mit Schulleitungen (Erhebungszeitpunkt t1) und Lehrpersonengruppen (Erhebungszeitpunkte t1, t2 und t3), eingesetzt (STEBLER et al. 2018). [14]

An der Studie waren neben der Projektleitung zahlreiche wissenschaftliche Mitarbeiter_innen, Doktorand_innen, Student_innen und Hilfskräfte beteiligt. Dies führte im Querschnitt zu arbeitsteiliger Kooperation und im Längsschnitt zu einer beträchtlichen personellen Fluktuation. Zum ersten Erhebungszeitpunkt waren beispielsweise drei Arbeitsgruppen an der Studie an den Erhebungs- und Auswertungsverfahren beteiligt. So wurden die Interviews mit den Schulleitungen und jene mit den Lehrpersonengruppen von zwei unterschiedlichen Teams vorbereitet und durchgeführt. Die beiden Leitfäden für die halbstrukturierten Interviews wurden jeweils von Mitgliedern der Forschungsgruppe in Zusammenarbeit mit Studierenden im Rahmen von Forschungsseminaren und -praktika entwickelt. Dabei wurden der theoretische Hintergrund der Forschungsfragen und der Forschungsstand aufgearbeitet, der Leitfaden konzipiert und die Interviewfragen formuliert. Für die Konzeption des Leitfadens waren einerseits die Forschungsfragen zentral, andererseits aber auch methodische Überlegungen – die Daten sollten induktiv ausgewertet werden können. Der Durchführung der Interviews ging eine entsprechende Schulung voraus (HELFFERICH 2011). Die Transkription der Interviews wurde von der dritten Arbeitsgruppe, bestehend aus Hilfskräften, übernommen. Am Ende der ersten Interviewphase lagen je zwölf Transkripte von Interviews mit Schulleitungen sowie mit Lehrpersonengruppen vor. [15]

3.2 Qualitätssicherung der Interviewdaten: Herausforderungen und Lösungen

Im Zuge der Datenaufbereitung wurde zunächst eine Interviewdatenbank eingerichtet. Dazu wurden alle Transkripte in das Programm MAXQDA eingelesen. Die Datenbank wurde über die gesamte Studiendauer zentral verwaltet: Alle Arbeitsschritte und Ergebnisse wurden gemäß internen Richtlinien für die Datenverwendung, Dokumentation und Speicherung registriert und protokolliert. Die Rechte, an den Projektdaten (MAXQDA-Datenbank und Transkripte) Änderungen vorzunehmen, wurden auf wenige Personen beschränkt. [16]

Die Auswertungen erfolgten mittels einer inhaltlich strukturierenden qualitativen Inhaltsanalyse (KUCKARTZ 2018). Begonnen wurde mit den Interviewtranskripten der Schulleitungen. Zuerst wurde eine Basiscodierung (*initial coding, erster Codierprozess*; BARBOUR 2014, S.499) durchgeführt. Dieses Vorgehen, bei dem sämtliche Textpassagen inhaltlich weit gefassten Hauptkategorien zugeordnet werden, erlaubt eine ökonomische Strukturierung großer Datenmengen (KUCKARTZ 2018). Die Hauptkategorien für die Basiscodierung wurden von zwei Mitgliedern des Projektteams aus den Themenblöcken⁴ des Interviews abgeleitet und bei der Datenanalyse induktiv ergänzt (BARBOUR 2014; KELLE & LAURIE 1995). Die Codierregeln wurden als Verfahrensregeln, die das Zuordnen der Textpassagen zu den Kategorien reglementieren, festgehalten (KUCKARTZ, DRESING, RÄDIKER & STEFER 2008). Die Codierung war insofern niederschwellig angelegt, als nur sechs breit gefasste Kategorien vorlagen, denen jeweils relativ große Datenblöcke (2000 bis 3000 Wörter) zugeordnet werden mussten, d.h., es war noch keine inhaltlich oder verfahrenstechnisch anspruchsvolle Strukturierung anhand von Subkategorien erforderlich. [17]

Bei der Erprobung codierten die beiden Teammitglieder zuerst unabhängig voneinander das gleiche Transkript. Nach dem Codieren wurden die Ergebnisse mittels der MAXQDA-Funktion zur Analyse der Intercoder-Übereinstimmung verglichen und diskursiv ein Konsens bezüglich Kategoriensystem und Codierregeln gesucht (HOPF & SCHMIDT 1993). Diesen Schritt wiederholten die beiden Codierer_innen mit drei weiteren Interviews. Beim dritten Interview war die Entwicklung der Codierregeln und der Kategorienbeschreibungen so weit fortgeschritten, dass dem Forschungsteam eine Prüfung der Intercoder-Übereinstimmung geboten erschien. Zur Wertung als Übereinstimmung war festgelegt worden, dass zwei Codierungen auf Segmentebene zu 90% übereinstimmen müssen (KUCKARTZ 2018). [18]

Es zeigte sich mit Blick auf alle sechs Hauptkategorien eine durchschnittliche Übereinstimmung von 81%. Bei zwei Kategorien erreichte die Übereinstimmung allerdings nur 66%⁵. Dieser niedrige Wert ist im Kontext einer Fehlertoleranz von 90% zu interpretieren, d.h. nur 66% der Kategoriengrenzen wurden bei einer Übereinstimmung von 90% von zwei unabhängigen Codierenden identisch gesetzt. Auch wenn eine durchschnittliche Übereinstimmung von 81% grundsätzlich akzeptabel ist, wurde dieser Wert im Kontext der breit angelegten Regeln und eines diskursiven, methodenspezifischen Vorgehens, welches geschulte Teammitglieder durchführten, als unbefriedigend beurteilt. Für diese Einschätzung waren noch zwei weitere Gründe wichtig: Erstens sollte das

4 Themenblöcke: 1. Schulleitung, 2. Schulkonzept und Schüler_innenschaft, 3. Unterricht, 4. Zusammenarbeit der Lehrpersonen, 5. Schul- und Unterrichtsentwicklung sowie 6. Ergänzungen

5 In vier Hauptkategorien (Schulleitung, Schulkonzept und Schüler_innenschaft, Unterricht, relevante Ergänzungen) stimmten die Zuordnungen der beiden Codierenden überein. Die Zuordnung der Kategorien "Zusammenarbeit Lehrpersonen" und "Schul- und Unterrichtsentwicklung" wich je zweimal voneinander ab. Daraus ergab sich eine Übereinstimmung von 66%. Über sechs Kategorien wurde in der Basiscodierung insgesamt eine Übereinstimmung der Codierungen von 81% erreicht.

Verfahren der Basiscodierung so robust sein, dass es von mehreren Codierer_innen in sehr ähnlicher Weise auf die vorliegenden sowie auf die zukünftigen Transkripte (t2 und t3) angewendet werden konnte. Dabei ging es nicht um die Objektivität und Reliabilität der Auswertung, sondern darum, die Daten zu den projektspezifischen Forschungsfragen im Längsschnitt so aufzubereiten, dass die inhaltliche Kohärenz über die drei Erhebungszeitpunkte gewährleistet und nachvollziehbar war (z.B. wurde bei der Beschreibung von Kategorien ein projektspezifisches Glossar angewendet). Zweitens bildeten die Basiscodierungen die Ausgangslage für weitere qualitativ inhaltsanalytische Auswertungen zu unterschiedlichen Themen. Im *zweiten Codierprozess* wurden die Hauptkategorien theoriegeleitet und induktiv erweitert. Diese deduktiv-induktive Erarbeitung von Subkategorien pro Hauptkategorie erfolgte im Rahmen mehrerer Forschungs- und Qualifikationsarbeiten und bezogen auf unterschiedliche Fragestellungen. Deshalb war es für die Forschungsgruppe wichtig, dass die Basiscodierung als Ausgangslage für weitere qualitativ inhaltsanalytische Auswertungen eine höhere Inter-coder-Übereinstimmung aufwies. Die berichteten Ergebnisse und Anforderungen führten zu folgenden Überlegungen mit Bezug auf die weiteren Datenerhebungen im Längsschnittdesign:

- *Gliederung des Leitfadens*: Die Themenblöcke des Interviewleitfadens zum Zeitpunkt t1 wurden theoriegeleitet und anhand der Forschungsfragen sowie einer Analyse von Dokumenten zu den Fallschulen der Studie konzipiert. Im ersten Codierprozess erwies sich die Gliederung des Leitfadens (z.B. die Reihenfolge der Frageblöcke) für die Codierung als herausfordernd (z.B. hätte der zweite Themenblock "Schulkonzept und Schüler_innenschaft" thematisch getrennt werden können). Weiter hätten inhaltliche Überschneidungen im Leitfaden gezielter vermieden werden können. So wurden z.B. im ersten Themenblock "Schulleitung" explizit Fragen zu den Aufgabenbereichen der Schulleitungen gestellt. Diese Aufgabenbereiche wurden teils auch in anderen Themenblöcken angesprochen. Bei relativ stark strukturierten Interviews, wie in unserem Praxisbeispiel, ist im ersten Codierprozess eine hohe Inter-coder-Übereinstimmung umso wahrscheinlicher, je genauer sich die Interviewenden an die Gliederung des Leitfadens halten und je deutlicher sie Themensprünge artikulieren und dokumentieren.
- *Gesprächsführung*: Die Interviews wurden von mehreren Interviewenden durchgeführt, wobei sich bezüglich der Gesprächsführung gewisse Unterschiede zeigten. Einzelne Interviews wurden eher offen und problemzentriert geführt (WITZEL & REITER 2012). Dabei wurde auf einige Themenblöcke sehr intensiv eingegangen, teilweise schweiften die Interviewten auch ab. Entsprechend weniger Zeit blieb für die anderen Themenblöcke, wodurch einzelne Interviewfragen nicht vertieft werden konnten. Die Interviewführung trug somit zur Varianz der Datenmenge pro Themenblock bei.
- *Anwendung des Leitfadens, thematische Elaboration*: Die Interviews wurden leitfadengestützt durchgeführt, weshalb im deduktiven Teil der qualitativen

Inhaltsanalyse die Leitfrage des Themenblocks als zentrales Kriterium für die Zuordnung einer Textstelle zu einer Hauptkategorie definiert wurde. Induktiv wurde die Kategorie anhand der angesprochenen Gesichtspunkte erweitert und beschrieben. In den Interviewsituationen traten spontane Themenwechsel und neue Gesichtspunkte auf, was nach FLICK (2014) legitim ist. Das Interview wurde zum Fachgespräch. Durch den flexiblen Einsatz des Interviewleitfadens konnten die Gesprächspartner_innen ihre Geschichte erzählen (GERGEN 2014; JOSSELSOON 2013). Spontane Themenwechsel oder Nebenbemerkungen seitens der Interviewten als Merkmale authentischer Kommunikation können zu mehr Tiefe und Breite des Themenspektrums führen und sind in Anbetracht der grundlegenden Offenheit der qualitativen Inhaltsanalyse methodisch unbedenklich (FLICK 2012). Jedoch kam es teils auch bei den Interviewer_innen zu spontanen Modifikationen der Leitfragen. Daraus folgten elaborierte Ausführungen zu Teilaspekten der thematischen Frageblöcke, die bei der Basiscodierung nicht immer eindeutig zugeordnet werden konnten, was zu einer Verringerung der InterCoder-Übereinstimmung führte. [19]

Aufgrund der Erfahrungen bei der ersten Erhebung (t1) wurden für die zweite Erhebung (t2) folgende Richtlinien formuliert, welche dazu dienen sollten, die Erhebungs- und die Auswertungsprozesse besser aufeinander abzustimmen, um eine höhere InterCoder-Übereinstimmung zu erzielen. Die Interviewer_innen

- beteiligen sich an der Konstruktion und der Pilotierung des Leitfadens;
- vertiefen ihre Kenntnisse bzgl. der Methode des halbstrukturierten Leitfadeninterviews;
- erlernen die Auswertungsmethode (inklusive deren Anwendung in MAXQDA), auch wenn sie die Interviews danach nicht selbst auswerten; so erfahren sie, wie die Interviewführung die Codierung beeinflusst;
- pilotieren die Interviewführung mit dem Originalleitfaden;
- entwickeln gemeinsam verbindliche Regelungen zur Anwendung des Leitfadens, zum Zeitmanagement sowie zur sprachlichen Markierung von Themensprüngen während des Interviews und zum Protokollieren von Kontextinformationen und
- tauschen während der Erhebungsphase Erfahrungen aus. [20]

In der zweiten Erhebungsphase (t2) der Studie, in der sechs Interviewer_innen elf Leitfadeninterviews mit Gruppen von Lehrpersonen führten, wurden diese Richtlinien angewendet. Die Basiscodierung der zweiten Erhebungsphase wurde analog zur ersten Phase durchgeführt und die InterCoder-Übereinstimmung überprüft, welche bei 90% Übereinstimmung nun 100% betrug. [21]

Zusammenfassend lässt sich festhalten, dass die Datenqualität der ersten Erhebung den Analyseprozess insofern beeinflusste, als sie zu einer unbefriedigenden InterCoder-Übereinstimmung führte. Dies veranlasste uns, die Erhebung und Analyse der Daten der Erhebungsphasen t2 und t3 vorgängig

stärker aufeinander abzustimmen und die interindividuelle Varianz bei der Umsetzung durch eine Anpassung der Regelungen einzuschränken, um die Qualitätssicherung nicht nur innerhalb der einzelnen Erhebungsphasen, sondern auch phasenübergreifend zu gewährleisten. Da im Längsschnittdesign Fragen der Schul- und Unterrichtsentwicklung im Vordergrund standen, dienten die qualitativ inhaltsanalytischen Auswertungen der einzelnen Erhebungszeitpunkte als Ausgangslage für die Entwicklung des Interviewleitfadens des jeweils nächsten Erhebungszeitpunkts. Zudem waren verschiedene Teams in den Forschungsprozess involviert. Daher musste die Qualitätssicherung so organisiert werden, dass nachfolgende Teams nahtlos an die Arbeiten der vorangehenden Teams anschließen konnten. Das beschriebene Vorgehen, bei dem ein methodenspezifisches Kriterium (Intercoder-Übereinstimmung) und prozessorientierte Strategien der Qualitätssicherung (Richtlinien für Interviews und Codierprozess) integrativ angewendet sowie die Prozesse der Datenerhebung und Datenauswertung inhaltlich und verfahrenstechnisch aufeinander abgestimmt wurden, wirkte sich günstig auf den gesamten Forschungsprozess aus. [22]

3.3 Konsequenzen

Unter Rückbezug auf die Ausführungen zur Qualitätssicherung (Abschnitt 2) in Verbindung mit den geschilderten Herausforderungen und Lösungen im Praxisbeispiel formulieren wir folgende Erkenntnisse: [23]

In der qualitativen Forschung *spielen die Forscher_innen bei der Qualitätssicherung eine wichtige Rolle*: Sie können die Qualität der Daten direkt beeinflussen; im Praxisbeispiel u.a. durch die Interviewführung. Folglich wird hier mit Blick auf qualitative Längsschnittstudien in großen Forschungsprojekten für eine verstärkte Qualitätssicherung plädiert, die auf die Instruktion und Schulung der Forscher_innen abzielt (HARVEY & GREEN 2000). Im dargestellten Beispiel zeigte sich, dass durch die integrativ angewendete Form der Qualitätssicherung die Qualität im gesamten Forschungsprozess gesteigert werden konnte. In der ersten Erhebungsphase resultierte aus dem Spielraum bei der Umsetzung der Leitfadeninterviews eine eher heterogene Datenmenge, was im Analyseprozess zu einer unbefriedigenden Intercoder-Übereinstimmung führte. Aufgrund dieses Befundes plädieren wir dafür, bei großen Forschungsprojekten Datenerhebung und -analyse nicht nur innerhalb der Erhebungsphasen, sondern auch im Längsschnitt vorgängig möglichst gut aufeinander abzustimmen und bei der Umsetzung die interindividuelle Varianz in und zwischen den Forschungsgruppen durch verbindliche Regelungen zu minimieren. Im Beispiel konnten durch entsprechende Maßnahmen der Qualitätssicherung die Deutungen und Vorgehensweisen in und zwischen den Forschungsgruppen vereinheitlicht werden. Dies wirkte sich einerseits methodenspezifisch günstig auf die Intercoder-Übereinstimmung der nächsten qualitativ inhaltsanalytischen Datenauswertung aus. Andererseits trug es zur formalen und inhaltlichen Kohärenz der aufeinanderfolgenden Erhebungen im Projektverlauf bei. Das heißt, die Themenblöcke wiesen eine ähnliche formale Struktur auf; gleichzeitig konnten

die Forscher_innen mit den thematischen Leitfragen inhaltlich auf die Aussagen der Interviewten im letzten Erhebungszeitpunkt eingehen. [24]

Die Forderung, *Erhebungs- und Auswertungsverfahren inhaltlich und verfahrensmäßig abzustimmen*, bezeichnen FLICK (2014) und STEINKE (1999, 2013) als Indikation. Wir übertrugen diese Vorgehensweise auf die computergestützte qualitative Inhaltsanalyse der Leitfadeninterviews in großen Forschungsgruppen und begründen mit der errechneten Inter-coder-Übereinstimmung die zentrale Bedeutung dieser Maßnahme für die Qualitätssicherung. Die Qualitätsfrage ist schließlich nicht nur für Datenerhebung und Datenauswertung relevant, sondern auch mit Bezug auf den Forschungskontext, das gesamte Forschungsdesign und die Publikation der Ergebnisse (BERGMAN & COXON 2005; HUBER 2001). [25]

Ein *einheitliches, gemeinsames Verfahrens- und Dokumentationssystem* ist zum einen für eine qualitativ hochwertige Arbeit innerhalb einer großen Forschungsgruppe unerlässlich, um die Qualität jeden Schrittes absichern und überprüfen zu können. Zum anderen erleichtert es die Legitimation der Forschung nach außen, weil beispielsweise durch die Verfahrensbeschreibung die intersubjektive Nachvollziehbarkeit (BERGMAN et al. 2010; STEINKE 2013) und Glaubwürdigkeit (ELO et al. 2014; FLICK 2018) gewährleistet werden. Ferner kann mit einem solchen System die Forderung nach reflektierter Subjektivität (STEINKE 1999) erfüllt werden. Je genauer die Dokumentation erfolgt, desto präziser können Grenzen zwischen den Forscher_innen und dem Feld gezogen und beschrieben werden. [26]

Ein *gemeinsam entwickeltes und geteiltes Verständnis von Instrumenten und Verfahren* ist in großen Forschungsgruppen, in denen auch Studierende bei der Erhebung und Auswertung mitarbeiten, für die Qualitätssicherung unabdingbar. KUCKARTZ (2018) verweist in diesem Zusammenhang auf Codierungsschulungen. In die hier vorgeschlagenen Maßnahmen zur Qualitätssicherung ist darüber hinaus die Erhebungsmethodik in das Schulungsprogramm mit aufgenommen worden. Zudem wird das Zusammenwirken von Datenerhebung und Datenauswertung explizit thematisiert. Auch sollten Ausführungen zum Qualitätsmanagement in Projektdesigns (etwa bei Bewerbungen um Drittmittel) unserer Ansicht nach explizit ausweisen, wie Forscher_innen für die Anwendung des gewählten methodologischen Instrumentariums qualifiziert werden. Beispielsweise macht es einen großen Unterschied, ob für ein multimethodisches, längsschnittliches Forschungsdesign über mehrere Erhebungszeitpunkte zehn oder 60 qualitative Interviews vorgesehen sind und diese von einem Kernteam oder von unterschiedlichen wissenschaftlichen Hilfskräften durchgeführt, aufbereitet und ausgewertet werden. [27]

Forschungssoftware gehört mittlerweile zum Forschungsalltag. Werden diese Computerprogramme nicht erst für die Auswertungen, sondern bereits in der Vorbereitung der Datenerhebungen eingesetzt, können sie der Qualitätssicherung in großen Forschungsgruppen dienen. Aufgrund der bisher

beschriebenen Erfahrungen gehen wir davon aus, dass die Ausbildungseinheiten mit MAXQDA bei der Interviewschulung zu einem gemeinsam geteilten Qualitätsverständnis beitrugen, was im Beispiel zu einer Verringerung der interindividuellen Varianz bei der Datenerhebung und schließlich zu einer höheren InterCoder-Übereinstimmung führte. Auf diese Weise konnten Datenerhebung und Datenanalyse formal und inhaltlich aufeinander bezogen werden. Wie oben dargelegt, wurden die Interviewer_innen nicht nur in der Interviewführung, sondern auch in der computergestützten qualitativen Inhaltsanalyse geschult (auch dann, wenn sie nicht an der Datenauswertung beteiligt waren). Die Schulung erfolgte u.a. in einem Methodenseminar, in dem die Studierenden den Interviewleitfaden konzipierten und die Interviews durchführten. Dank der Schulung konnten die Interviewer_innen die Relevanz ihrer Interviewführung für die Datenauswertung besser nachvollziehen. Im Beispiel wurden die Daten für Auswertungen verschiedener Fragestellungen genutzt. Über die zentral eingerichtete und verwaltete Datenbank hatten die Forscher_innen einen effizienten Zugang zu den Daten. Aufgrund der BasisCodierung, bei der im Längsschnitt teils identische Codes verwendet wurden, waren die relevanten Daten für verlaufsbezogene Auswertungen leicht verfügbar. Generell erleichterte die Software die Auswahl von Daten für weiterführende Analysen. Software ist somit ein wertvolles Instrument, um den Forschungsprozess effizienter und transparenter zu machen (FLICK 2007a), was für eine breite Nutzung der Daten für qualitätsvolle Primär- und Sekundäranalysen unerlässlich ist (KRÜGER 2000; MEDJEDOVIĆ & WITZEL 2008). Damit kann auch die für multimethodische Studien charakteristische Forderung, die Daten der qualitativen und der quantitativen Studienteile bei den Auswertungen aufeinander zu beziehen (KNAPPERTSBUSCH 2017), eingelöst werden. [28]

4. Maßnahmen zur Qualitätssicherung für große Forschungsgruppen

Aus den bisherigen Erörterungen leiten wir im Folgenden vier Maßnahmen zur Qualitätssicherung in qualitativen Studienteilen großer multimethodischer Längsschnittstudien ab (Abbildung 1):

1. *Erhebungs- und Auswertungsmethoden aufeinander abstimmen:* Bei der BasisCodierung (BARBOUR 2014) erscheint uns eine hohe Übereinstimmung der Codierenden umso wahrscheinlicher, je stringenter die Regeln zur Interviewführung von den Interviewerinnen und Interviewern eingehalten werden (z.B. je deutlicher sie Themensprünge artikulieren). Diese Maßnahme der Qualitätssicherung trägt aus unserer Sicht dazu dabei, qualitative (Teil-)Studien verfahrenstechnisch und thematisch optimal fokussiert zu realisieren, was besonders in Längsschnittstudien unerlässlich ist und in multimethodischen Studien wesentlich dazu beiträgt, generalisierbare Ergebnisse mit hinreichender externer Validität zu erhalten.
2. *Ein einheitliches Verfahrens- und Dokumentationssystem festlegen:* Die Forschungsprozesse sollten nach schriftlichen Richtlinien zu den Arbeitsschritten, zur Dokumentation der Erhebungen (z.B. Inhalt und Umfang

der Protokollierung von Kontextinformationen, zu Speicher- und Sicherungssystemen, zu internen Bezeichnungen und zu Korrekturvorgängen ausgeführt werden. Verbindliche Richtlinien tragen insofern zur Qualitätssicherung bei, als sie die interindividuelle Varianz im gesamten Forschungsprozess reduzieren und eine produktive Nutzung von Daten und Ergebnissen durch unterschiedliche Personen in multimethodischen Quer- und Längsschnittstudien wesentlich erleichtern können.

3. *Forscher_innen mit Instrumenten und Verfahren vertraut machen:* Forscher_innen sollten innerhalb des Forschungsvorhabens bezüglich Methoden der Datenerhebung und Datenauswertung geschult und für deren wiederholte Anwendung in Längsschnittstudien sensibilisiert werden (u.a. Pilotieren des Leitfadens, Üben der Interviewsituation, Kennenlernen der Forschungssoftware, Anwenden der Analysemethode und -instrumente sowie der Codierungsarbeit in Schulungen). Dadurch können Verfahrensfehler minimiert und Abläufe vereinheitlicht werden. Besonders wichtig ist in diesem Zusammenhang, dass die Forscher_innen auch grundlegende Kenntnisse studienrelevanter Instrumente und Verfahren erwerben, die sie selbst nicht anwenden müssen. Mit solchen Kenntnissen kann sowohl das eigene Forschungshandeln als auch die Qualität von Prozessen und Produkten günstig beeinflusst werden.
4. *Qualitätskontrolle einzelner methodischer Schritte und Verfahren:* Bei der qualitativ inhaltsanalytischen Datenauswertung können Qualitätskontrollen z.B. mittels Interocoder-Übereinstimmung (KUCKARTZ 2018; MAYRING 2015) und konsensuellem Codieren (HOPF & SCHMIDT 1993) erfolgen. Bei anderen Zugängen in der qualitativen Forschung sind entsprechende methoden- und verfahrensspezifische Kriterien der Qualitätssicherung heranzuziehen.

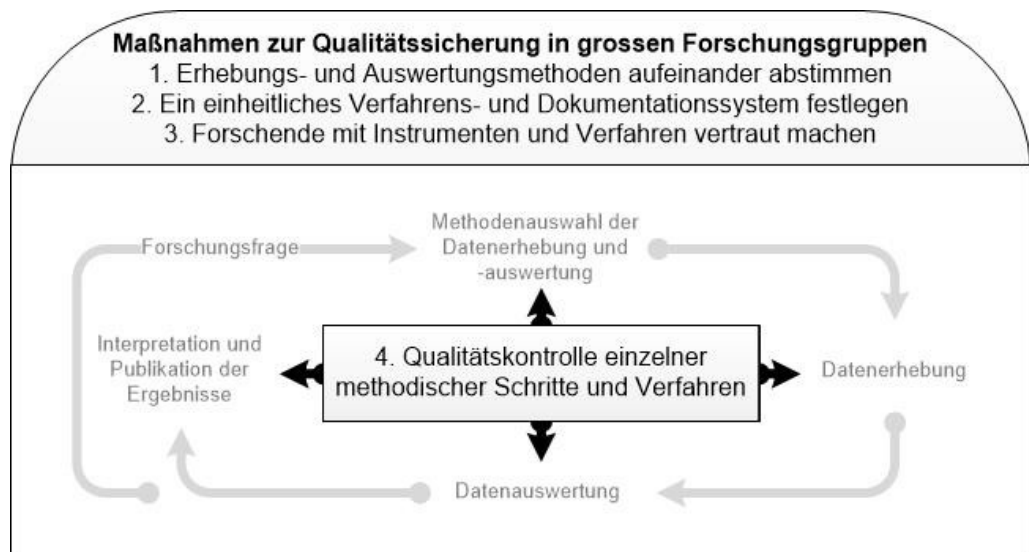


Abbildung 1. Maßnahmen zur Qualitätssicherung in großen Forschungsgruppen [29]

An die Durchführung qualitativer, längsschnittlich konzipierter Teilstudien in großen Forschungsprojekten werden bezüglich Qualitätssicherung in mehrfacher

Hinsicht hohe Anforderungen gestellt. Aufgrund der eigenen Erfahrungen im Praxisbeispiel scheint es uns sinnvoll, dass die beschriebenen Maßnahmen (Kriterien und Strategien) erstens in jeder einzelnen Erhebungsphase angewendet, reflektiert und angepasst werden sowie zweitens die Qualitätssicherung im Querschnitt so geplant und umgesetzt wird, dass methodische Stabilität und inhaltliche Kohärenz im Längsschnitt garantiert werden. Schließlich wäre das Vorgehen bei der Qualitätssicherung so zu wählen, dass sich Daten und Ergebnisse aus qualitativen und quantitativen Teilstudien ohne inhaltliche und/oder methodische Vorbehalte aufeinander beziehen lassen. Wenn große, multimethodische Designs den Königsweg in der Sozialforschung bilden sollen (FUCHS, LAMNEK & LUEDTKE 2001), können studienspezifisch und kompetent orchestrierte Maßnahmen der Qualitätssicherung in den qualitativen Teilstudien als Leitplanken dieses Weges betrachtet werden. [30]

Danksagung

Wir danken den zahlreichen Mitwirkenden im Projekt perLen – Studierenden, wissenschaftlichen Hilfskräften und Mitarbeitenden – für die lehrreichen Erfahrungen an einer methodischen Großbaustelle. Unser besonderer Dank gilt den beiden Gutachter_innen dieses Beitrags für ihre gehaltvollen und anregenden Rückmeldungen.

Literatur

- Asbrand, Barbara; Pfaff, Nicole & [Bohnsack, Ralf](#) (2013). Rekonstruktive Längsschnittforschung in ausgewählten Gegenstandsfeldern der Bildungsforschung. *Zeitschrift für Qualitative Forschung*, 14(1), 3-12, <https://doi.org/10.3224/zqf.v14i1.15449> [Zugriff: 19. Juli 2019].
- Barbour, Rosaline (2014). Quality of data analysis. In [Uwe Flick](#) (Hrsg.), *The Sage handbook of qualitative data analysis* (S.496-509). Thousand Oaks, CA: Sage.
- Bergman, Manfred Max & Coxon, Anthony P. M. (2005). The quality in qualitative methods. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(2), Art. 34, <http://dx.doi.org/10.17169/fqs-6.2.457> [Zugriff: 19. Juli 2019].
- Bergman, Manfred Max; [Eberle, Thomas S.](#); Flick, Uwe; Förster, Till; Horber, Eugène; Maeder, Christoph; Mottier, Véronique; Nadai, Eva; Rolshoven, Johanna; Seale, Clive & Widmer, Jean (2010). *Methoden qualitativer Sozialforschung – Manifest zur Bedeutung, Qualitätsbeurteilung und Lehre der Methoden qualitativer Sozialforschung*. Bern: SAGW, https://www.alexandria.unisg.ch/61300/1/Bergman_Manfred_Eberle_Thomas_2010_Manifest%20zur%20qualitativen%20Sozialforschung.pdf [Zugriff: 19. Juli 2019].
- Bohl, Thorsten & Wacker, Albrecht (Hrsg.) (2016). *Die Einführung der Gemeinschaftsschule in Baden-Württemberg. Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem)*. Münster: Waxmann.
- Bohnsack, Ralf & Krüger, Heinz-Hermann (2005). Qualität qualitativer Forschung. Einführung in den Themenschwerpunkt. *Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung*, 6(2), 185-190, <https://www.ssoar.info/ssoar/handle/document/27826> [Zugriff: 6. September 2019].
- Bourdon, Sylvain (2002). The integration of qualitative data analysis software in research strategies. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 11, <http://dx.doi.org/10.17169/fqs-3.2.850> [Zugriff: 19. Juli 2019].
- Braun, Annika; Weiss, Sabine & Kiel, Ewald (2018). Interkulturelle Schulentwicklung an Grundschulen. Die Perspektiven von Lehrer/innen und Schulleiter/innen. *Zeitschrift für Bildungsforschung*, 8, 121-134.
- Bryman, Alan (2006). Integrating quantitative and qualitative research. How is it done?. *Qualitative Research*, 6, 97-113.

[Denzin, Norman K.](#) & Lincoln, Yvonna S. (Hrsg.) (2005). *The Sage handbook of qualitative research* (2. Aufl.). Thousand Oaks, CA: Sage.

Döring, Nicola & Bortz, Jürgen (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Berlin: Springer.

Elo, Satu; Kääriäinen, Maria; Kanste, Outi; Pölkki, Tarja; Utriainen, Kati & Kyngäs, Helvi (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open*, 4(1), 121-135, <https://doi.org/10.1177/2158244014522633> [Zugriff: 19. Juli 2019].

Evers, Jeanine; [Mruck, Katja](#); Silver, Christina & Peeters, Bart in cooperation with di Gregorio, Silvana & Tagg, Clare (Hrsg.) (2011). The KWALON experiment: Discussions on qualitative data analysis software by developers and users. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 12(1), <http://www.qualitative-research.net/index.php/fqs/issue/view/36> [Zugriff: 20. September 2019].

Flick, Uwe (2005). Standards, Kriterien, Strategien – Zur Diskussion über Qualität qualitativer Sozialforschung. *Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung*, 6(2), 191-210.

Flick, Uwe (2007a). Zur Qualität qualitativer Forschung. In Udo Kuckartz, Heiko Grunenberg & Thorsten Dresing (Hrsg.), *Qualitative Datenanalyse: computergestützt. Methodische Hintergründe und Beispiele aus der Forschungspraxis* (2. Aufl., S.187-209). Wiesbaden: VS Verlag für Sozialwissenschaften.

Flick, Uwe (2007b). *Managing quality in qualitative research*. London: Sage.

Flick, Uwe (2011). *Triangulation. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Flick, Uwe (2012). *Qualitative Sozialforschung. Eine Einführung* (5. Aufl.). Reinbek: Rohwolt.

Flick, Uwe (2014). Gütekriterien qualitativer Sozialforschung. In Nina Baur & Jörg Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S.411-423). Wiesbaden: Springer VS.

Flick, Uwe (2017). Gütekriterien qualitativer Forschung in der Psychologie. In [Günter Mey](#) & Katja Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S.1-17). Wiesbaden: Springer VS.

Flick, Uwe (2018). Gütekriterien. In Leila Akremi, Nina Baur, [Hubert Knoblauch](#) & Boris Traue (Hrsg.), *Handbuch Interpretativ forschen* (S.183-202). Weinheim: Beltz.

Ford, Kate; Oberski, Iddo & Higgins, Steve (2000). Computer-aided qualitative analysis of interview data: Some recommendations for collaborative working. *The Qualitative Report*, 4(3/4), 1-16, <https://nsuworks.nova.edu/tqr/vol4/iss3/1> [Zugriff: 21. August 2019].

Frost, David M. (2014). Markers of quality and best practices in qualitative inquiry: Introduction to the special section. *Journal of Qualitative Research*, 1(1), 47-48.

Fuchs, Marek; [Lamnek, Siegfried](#) & Luedtke, Jens (2001). *Tatort Schule: Gewalt an Schulen 1994-1999*. Opladen: Leske + Budrich.

[Gergen, Kenneth J.](#) (2014). Pursuing excellence in qualitative inquiry. *Qualitative Psychology*, 1(1), 49-60, https://www.taosinstitute.net/Websites/taos/images/DegreeStudentResources/Pursuing_excellence.pdf [Zugriff: 6. September 2019].

[Gibbs, Graham R.](#); Friese, Susanne & Mangabeira, Wilma C. (2002). Technikeinsatz im qualitativen Forschungsprozess. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 8, <http://dx.doi.org/10.17169/fqs-3.2.847> [Zugriff: 19. Juli 2019].

Häbig, Julia (2018). *Lernentwicklungsgespräche aus der Sicht von Schülerinnen und Schülern*. Wiesbaden: Springer VS.

Harvey, Lee & Green, Diana (2000). Qualität definieren. Fünf unterschiedliche Ansätze. *Zeitschrift für Pädagogik*, 41(Beiheft), 17-40, urn:nbn:de:0111-opus-84835 [Zugriff: 6. September 2019].

Helferich, Cornelia (2011). *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews* (4. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Helsper, Werner (2016). Wird die Pluralität in der Erziehungswissenschaft aufgekündigt?. *Zeitschrift für Erziehungswissenschaft*, 19(1), 89-105.

Helsper, Werner; Kelle, Helga & Koller, Hans-Christoph (2016). Qualitätskriterien der Begutachtung qualitativer Forschungsvorhaben in der Erziehungswissenschaft. Ergebnisse eines DFG-Roundtable. *Zeitschrift für Pädagogik*, 62(5), 738-748, https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissen

[schaften/geisteswissenschaften/ergebnisse_roundtable_qualitaetskriterien_forschungsvorhaben_erziehungswissenschaft.pdf](#) [Zugriff: 19. Juli 2019].

Hopf, Christel & Schmidt, Christiane (1993). *Zum Verhältnis von innerfamiliären sozialen Erfahrungen, Persönlichkeitsentwicklung und politischen Orientierungen: Dokumentation und Erörterung des methodischen Vorgehens in einer Studie zu diesem Thema*, <http://nbn-resolving.de/urn:nbn:de:0168-ssaar-456148> [Zugriff: 19. Juli 2019].

Huber, Andreas (2001). Die Angst des Wissenschaftlers vor der Ästhetik. Zu Jo Reichertz: Zur Gültigkeit von Qualitativer Sozialforschung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 2(2), Art. 1, <http://dx.doi.org/10.17169/fqs-2.2.961> [Zugriff: 19. Juli 2019].

Josselson, Ruthellen (2013). *Interviewing for qualitative inquiry: A relational approach*. New York, NY: Guilford.

Kelle, Udo & Laurie, Heather (1995). Computer use in qualitative research and issues of validity. In Udo Kelle, Gerald Prein & Katherine Bird (Hrsg.), *Computer-aided qualitative data analysis. Theory, methods and practice* (S.19-28). London: Sage.

Kiener, Urs & Schanne, Michael (2001). Kontextualisierung, Autorität, Kommunikation. Ein Beitrag zur FQS-Debatte über Qualitätskriterien in der interpretativen Sozialforschung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 2(2), Art. 2, <http://dx.doi.org/10.17169/fqs-2.2.962> [Zugriff: 19. Juli 2019].

Knappertsbusch, Felix (2017). Ökologische Validität durch Mixed-Methods-Designs. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2, Suppl.), 337-360.

Krippendorff, Klaus (2013). *Content analysis. An introduction to its methodology*. Thousand Oaks, CA: Sage.

Krüger, Heinz-Hermann (2000). Stichwort: Qualitative Forschung in der Erziehungswissenschaft. *Zeitschrift für Erziehungswissenschaft*, 3(3), 323-342.

Krüger, Heinz-Hermann (2010). The importance of qualitative methods in the German educational science. In Ralf Bohnsack, Nicole Pfaff & Wivian Weller (Hrsg.), *Qualitative analysis and documentary method in international educational research* (S.53-74). Opladen: Barbara Budrich.

Kuckartz, Udo (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (4. Aufl.). Weinheim: Beltz.

Kuckartz, Udo; Dresing, Thorsten; Rädiker, Stefan & Stefer, Claus (2008). *Qualitative Evaluation* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Lamnek, Siegfried (2010). *Qualitative Sozialforschung* (5. Aufl.). Weinheim: Beltz.

Lamnek, Siegfried & Krell, Claudia (2016). *Qualitative Sozialforschung* (6., überarb. Aufl.). Weinheim: Beltz.

Li, Sarah & Seale, Clive (2007). Learning to do qualitative data analysis: An observational study of doctoral work. *Qualitative Health Research*, 17(10), 1442-1452.

Ludwig, Peter H. (2011). "Anything Goes – Quality Stays". Eine normative Positionierung zu Qualitätsstandards für begründungskontextuelle qualitative Forschung. *Empirische Pädagogik*, 25(1), 88-107.

Lüders, Christian (2013). Herausforderungen qualitativer Forschung. In Uwe Flick, Ernst von Kardorff & Ines Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (10. Aufl., S.632-642). Reinbek: Rowohlt.

Mayring, Philipp (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (12., überarb. Aufl.). Weinheim: Beltz.

Mayring, Philipp (2016). *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken* (6. Aufl.). Weinheim: Beltz.

Medjedović, Irena & Witzel, Andreas (2008). Secondary analysis of interviews: Using codes and theoretical concepts from the primary study. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(1), Art. 46, <http://dx.doi.org/10.17169/fqs-6.1.507> [Zugriff: 19. Juli 2019].

Pilz, Simone (2018). *Schulentwicklung als Antwort auf Heterogenität und Ungleichheit*. Wiesbaden: Springer VS.

Reinold, Martin (2016). *Lehrerfortbildungen zur Förderung prozessbezogener Kompetenzen. Eine Analyse der Effekte auf den Wirkungsebenen Akzeptanz und Überzeugungen*. Wiesbaden: Springer VS.

- [Schreier, Margrit](#) (2014). Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 15(1), Art. 18, <http://dx.doi.org/10.17169/fqs-15.1.2043> [Zugriff: 7. September 2019].
- Schütze, Fritz (2005). Eine sehr persönlich generalisierte Sicht auf qualitative Sozialforschung. *Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung*, 6(2), 211-248, <https://www.budrich-journals.de/index.php/zqf/article/view/2038/1666> [Zugriff: 19. Juli 2019].
- Seale, Clive (1999). *The quality of qualitative research*. Thousand Oaks, CA: Sage.
- Seale, Clive (2002). Quality issues in qualitative inquiry. In Paul Atkinson & Sara Delamont (Hrsg.), *Qualitative research methods* (S.98-275). Thousand Oaks, CA: Sage.
- Stamann, Christoph; Janssen, Markus & Schreier, Margrit (2016). Qualitative Inhaltsanalyse – Versuch einer Begriffsbestimmung und Systematisierung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(3), Art. 16, <http://dx.doi.org/10.17169/fqs-17.3.2581> [Zugriff: 19. Juli 2019].
- Stebler, Rita; Pauli, Christine & Reusser, Kurt (2018). Personalisiertes Lernen – Zur Analyse eines Bildungsschlagwortes und erste Ergebnisse aus der perLen-Studie. *Zeitschrift für Pädagogik*, 64(2), 159-178.
- Steinke, Ines (1999). *Kriterien qualitativer Forschung*. Weinheim: Juventa.
- Steinke, Ines (2013). Gütekriterien qualitativer Forschung. In Uwe Flick, Ernst von Kardorff & Ines Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (10. Aufl., S.319-331). Reinbek: Rowohlt.
- Strübing, Jörg (2018). *Qualitative Sozialforschung: Eine komprimierte Einführung* (2. Aufl.). Boston, MA: DeGruyter.
- Strübing, Jörg; Hirschauer, Stefan; Ayaß, Ruth; Krähnke, Uwe & Scheffer, Thomas (2018). Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß. *Zeitschrift für Soziologie*, 47(2), 83-100, <https://doi.org/10.1515/zfsoz-2018-1006> [Zugriff: 19. Juli 2019].
- Wertz, Frederick J. (2011). The qualitative revolution and psychology: Science, politics, and ethics. *The Humanistic Psychologist*, 39(2), 77-104.
- Witzel, Andreas & Reiter, Herwig (2012). *The problem-centred interview: Principles and practice*. Los Angeles, CA: Sage.

Zu den Autorinnen und Autoren

Katriina VASARIK STAUB ist wissenschaftliche Projektmitarbeiterin am Institut für Erziehungswissenschaft an der Universität Zürich. Sie forscht zur Schul- und Unterrichtsentwicklung mit Schwerpunkten innovative Lehr-Lernkonzepte, Unterrichtsstörungen und Beziehung zwischen Schule und Elternhaus.

Kontakt:

Katriina Vasarik Staub

Institut für Erziehungswissenschaft
Universität Zürich
Freiestrasse 36, 8032 Zürich

Tel.: +41 44 634 45 65

E-Mail: vasarik.staub@ife.uzh.ch

URL:

<https://www.ife.uzh.ch/de/research/ppd/mitarbeite2/vasarikkatriina.html>

Marco GALLE ist wissenschaftlicher Mitarbeiter an der Pädagogischen Hochschule Zürich. Seine Forschungsschwerpunkte sind Unterrichts- und Schulentwicklung, ko-konstruktive Kooperation, kooperative Lehrer_innenbildung und berufspraktische Ausbildung.

Kontakt:

Marco Galle

Abteilung Ausbildung
Pädagogische Hochschule Zürich
Lagerstrasse 2, 8090 Zürich

Tel.: +41 43 305 57 94

E-Mail: marco.galle@phzh.ch

URL: <https://phzh.ch/personen/marco.galle>

Rita STEBLER ist wissenschaftliche Mitarbeiterin am Lehrstuhl Pädagogisch-psychologische Lehr-Lernforschung und Didaktik am Institut für Erziehungswissenschaft (Universität Zürich). Ihre Forschungsschwerpunkte sind innovative Lehr-Lernkulturen, integrative Schulformen, mathematische Frühförderung und Kompetenzorientierung.

Kontakt:

Rita Stebler

Institut für Erziehungswissenschaft
Universität Zürich
Freiestrasse 36, 8032 Zürich

Tel.: +41 44 634 45 65

E-Mail: stebler@ife.uzh.ch

URL:

<https://www.ife.uzh.ch/de/research/lehrstuhlpraeetorius/mitarbeitende/steblerrita.html>

Kurt REUSSER ist emeritierter Professor für Pädagogische Psychologie und Didaktik. Seine Forschungsschwerpunkte sind u.a. Lehr-Lernpsychologie mit Schwerpunkt Mathematiklehren und -lernen, Lehren und Gestaltung von Lernumwelten, Allgemeine Didaktik und videobasierte Unterrichtsforschung, Lehrerinnen- und Lehrerbildung.

Kontakt:

Kurt Reusser

Institut für Erziehungswissenschaft
Universität Zürich
Freiestrasse 36, 8032 Zürich

Tel.: +41 44 634 45 65

E-Mail: reusser@ife.uzh.ch

URL:

<https://www.ife.uzh.ch/de/research/ppd/mitarbeitende2/reusserkurt.html>

Zitation

Vasarik Staub, Katriina; Galle, Marco; Stebler, Rita & Reusser, Kurt (2019). Qualitätssicherung bei qualitativ inhaltsanalytischen Verfahren in großen Forschungsgruppen: Herausforderungen und Möglichkeiten in der Forschungspraxis am Beispiel der perLen-Studie [30 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), Art. 33, <http://dx.doi.org/10.17169/fqs-20.3.3391>.